

In search for a common geographical base to Compare Statistics across Europe: The Tandem Project

- Philippe Guiblin, the Office of National Statistics, UK
- Marja Tammilehto-Luode, Statistics Finland
- Lars Backer, Statistics Sweden



The Tandem Consortium

Plan for the presentation

- Background and motivation
- The two parallel approaches
- Methods
- Tests with empirical data
- Results



The Tandem Consortium

A Feasibility study

- Made by the "Tandem Consortium" :Representatives from Statistics Finland, Statistics Sweden and the Office of National Statistics (UK)
- Funded by a grant of European Commission (85%) and the participating NSI's (15%)
- Final results October 2001
- Objective: to find alternative comparable building blocks/territorial division for European system of small area statistics
- This paper represents the views of the research team; it does not commit the NSIs for which they work



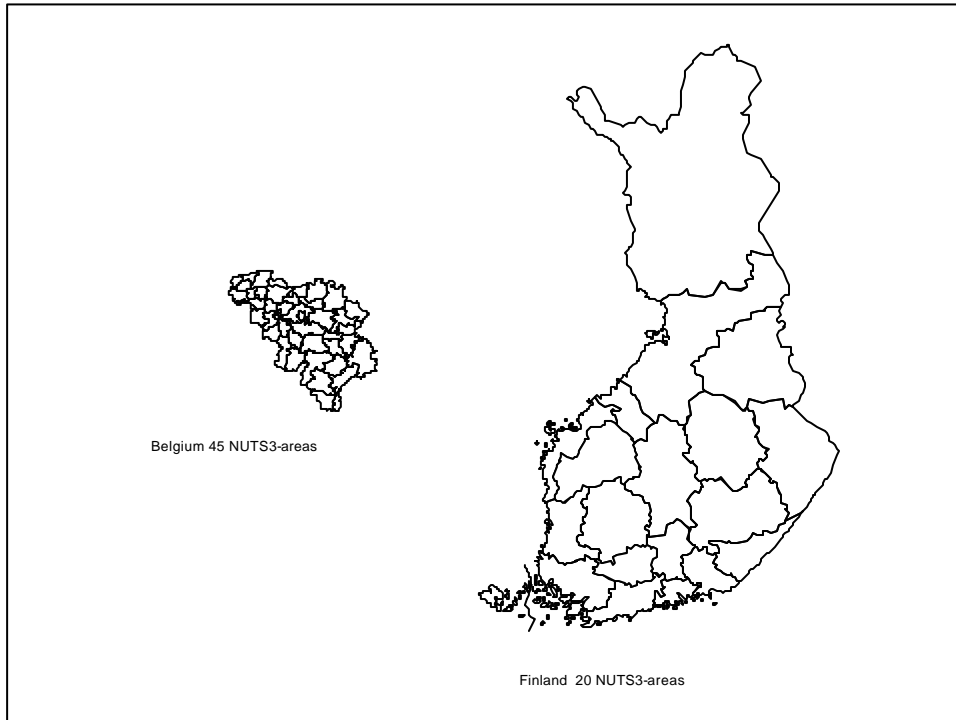
The Tandem Consortium

Point of departure: A need for more comparable territorial divisions for statistics

- to visualise data more effectively
- to combine and compare data on different spatial units
- to combine or compare data on different spatial scales
- to make better statistical/spatial analysis of data (to test spatial patterns and trends)



The Tandem Consortium



The study of two parallel approaches

- **Regular tessellation approach**
- Statistics Finland's responsibility
- study of methods to construct grid-based statistics
- tests of candidate methods with empirical data
- construction of a prototype
- tests on a prototype with a user case
- **Irregular tessellation approach**
- Office of National Statistics in UK's responsibility
- study of methods of zoning design
- tests with ZDES (Zone Design System) by Leeds University with empirical data
- tests on a prototype with a user case



The Tandem Consortium

Towards a system of Grid-based statistics

- **Methods**
 - Points to grid cells -need for best practices
 - Polygons to grid cells - different methods give different results - need for standardisation
- **Size of a grid cell** dependent on
 - quality of data
 - confidential grid cells
 - scale of the study
 - compatibility of different kind of source data
 - disk space, processing speed
- **Map projection**
 - rectangular
 - common geo-reference system - with a same origin



The Tandem Consortium

Towards a system of Polygon-based statistics

- Optimum grouping of the source areas
 - equal values of selected variable (population)
 - similar degree of heterogeneity (homogeneity)
 - weighted accessibility on certain input variable (shape)
- The automated zone design program (Openshaw 1977)
 - optimising an objective function
 - optimising under constraints
- Design functions
 - equality population zoning
 - shape design function
 - homogeneity design function
 - correlation -, distance -, spatial autocorrelation functions

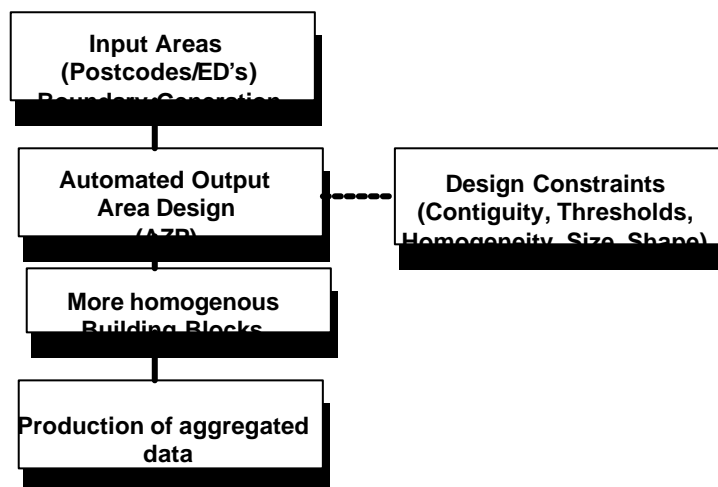


Polygons as building blocks

- Available software:
 - ZDES (Zone Design System) by Leeds University UK
 - SAGE (Spatial Analysis GIS Environment) by Sheffield University UK
- Tandem consortium tested ZDES



Building Blocks Production System



Building Blocks Production System

- Create small areas below Nuts 5 level
- Postcode-based and/or Enumeration District-based
- Must meet confidentiality thresholds
- Consistent throughout Europe



Tests:

Data from two countries, different types of geo-references, different types of variables

- **Grids**
 - tests of candidate algorithms to convert polygon-based data to grids
 - construction of prototype
 - visualisation of results
 - delineation of urban areas
- **Blobs**
 - tests of zone function with constraints of equal population and a single weighted shape
 - construction of prototypes
 - visualisation of results
 - delineation of urban areas



Test areas and data sets

- The **Helsinki data sets** cover the city of Helsinki and contain the census population counts. Two level of resolution are considered (one for each data set):
 - ?Population at post-code level;
 - ?Population on 1km x 1km grids.
- The **British data set** covers the county of South Glamorgan (NUTS3 containing the City of Cardiff and the Vale of South Glamorgan) and contain the 91 census population counts at Enumeration District (ED) level.



The Tandem Consortium

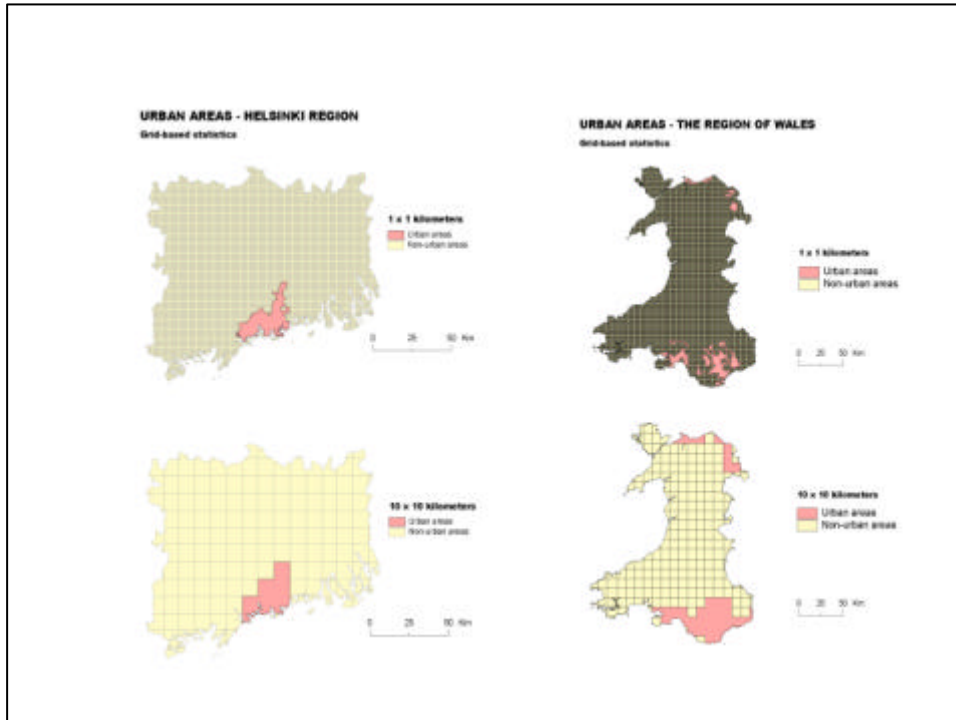


Building up a system of grids - Results

Tests on area-based data: from Polygons to grid cells

- Variables: population counts, population densities
- Algorithms:
 - Regiongrid and Pointgrids algorithms on counts
 - Polygrid algorithm on densities
- Grid size: 1 km x 1 km and 10 km x 10 km

- Regiongrid-algorithm maintains the original data structure (similar statistics to 'real grids')
- Larger grids keep the structure better than smaller
- Polygrid and Pointgrids algorithms overestimate data to small grids and underestimate data in larger grids

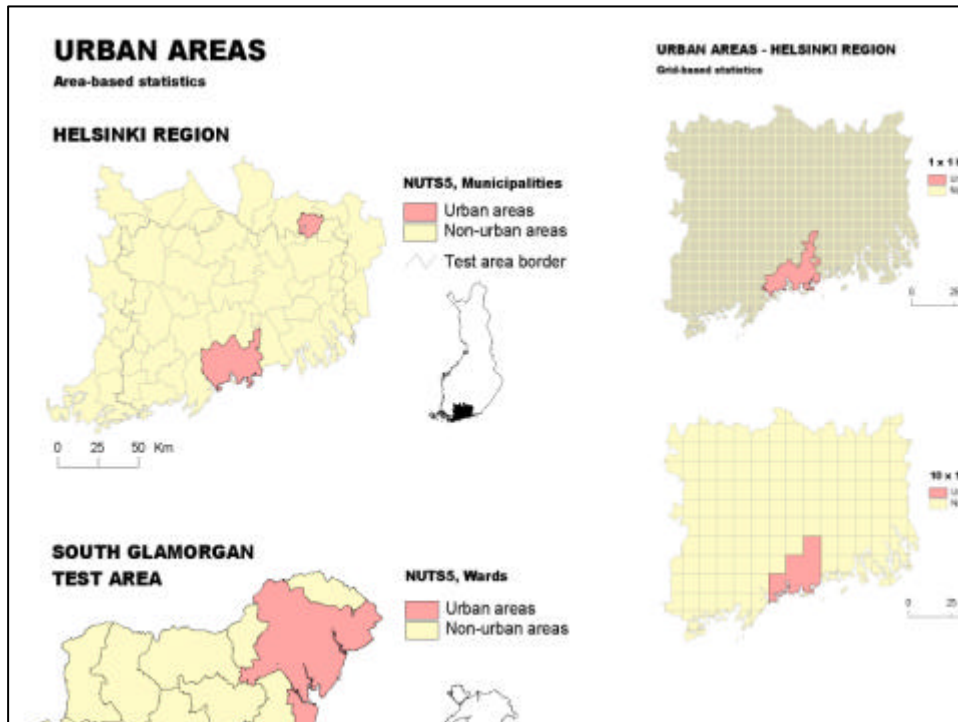


Building up a system of grids - Results

Tests on area-based data: Delineation of Urban areas

- The urban area defined using the Finnish test data relating to 1km x 1km grids is 45% smaller than that defined using the data on NUTS5 areas.
- On the other hand the urban area of the UK test area, South Glamorgan is a little bit larger by 1km x 1km building blocks than that defined using the data on NUTS5 areas.

This shows a big difference of areas of NUTS5 regions in the two different country.



Building up a system of blobs: AZP on British Enumeration District data

Data:

1991 British census population data at ED level for the county (NUTS3) of South Glamorgan (Wales)

Software:

Visual Visual Basic program, AZM, developed by David Martin (Southampton University)



Building up a system of blobs

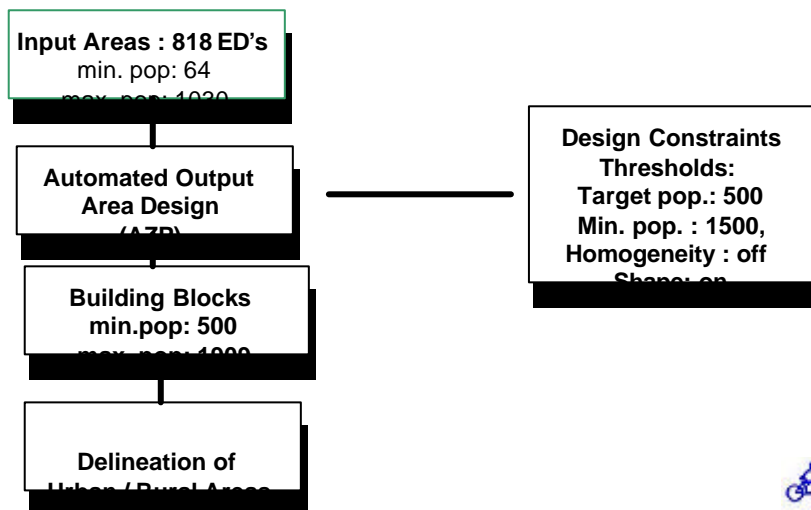
- Constraints:
- Only equal population zoning and a shape constraints (perimeter squared/area) were used.
- no other source of heterogeneity was considered such as geography and/or social class

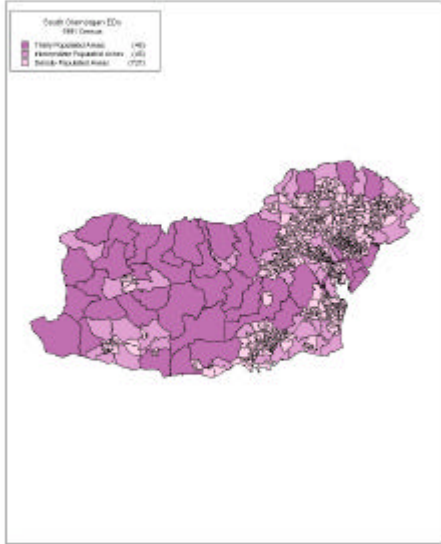
Results:

- Production of an 'optimised' boundary system.



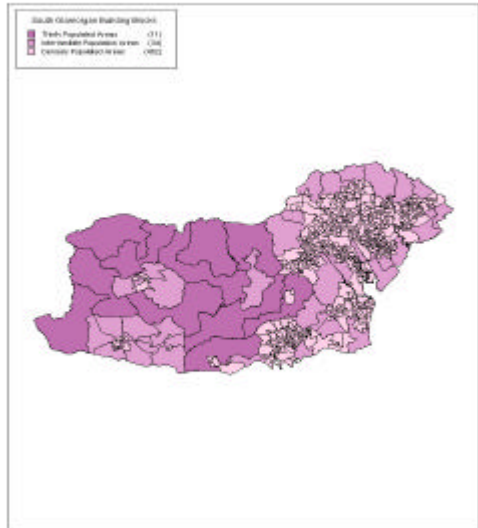
AZP Case Study: Cardiff Area





AZP on South Glamorgan Enumeration District

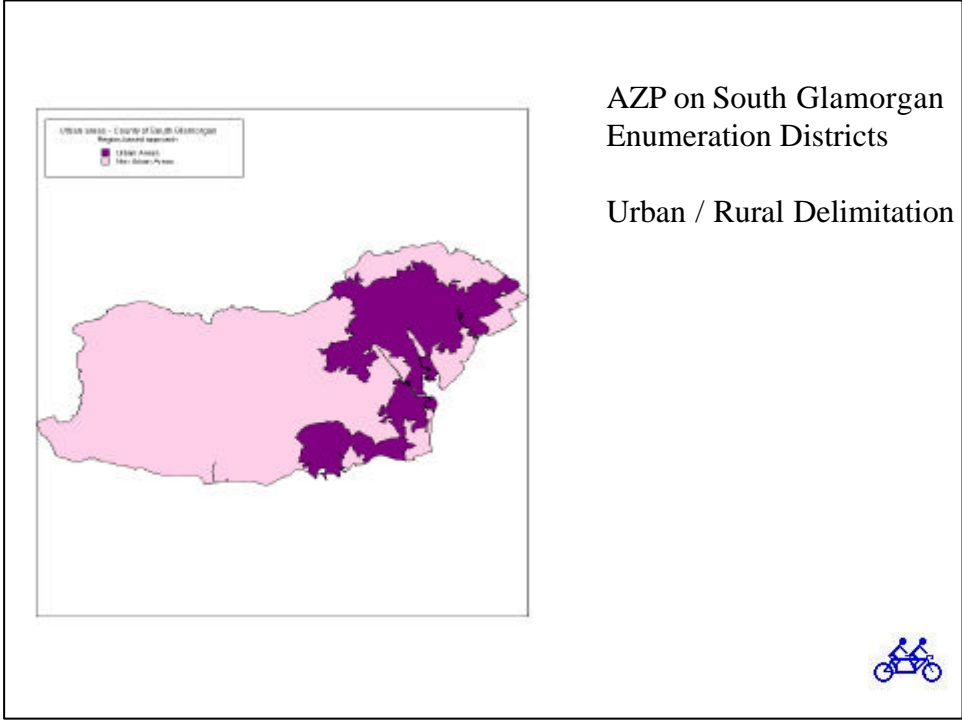
Input Areas



AZP on South Glamorgan Enumeration District

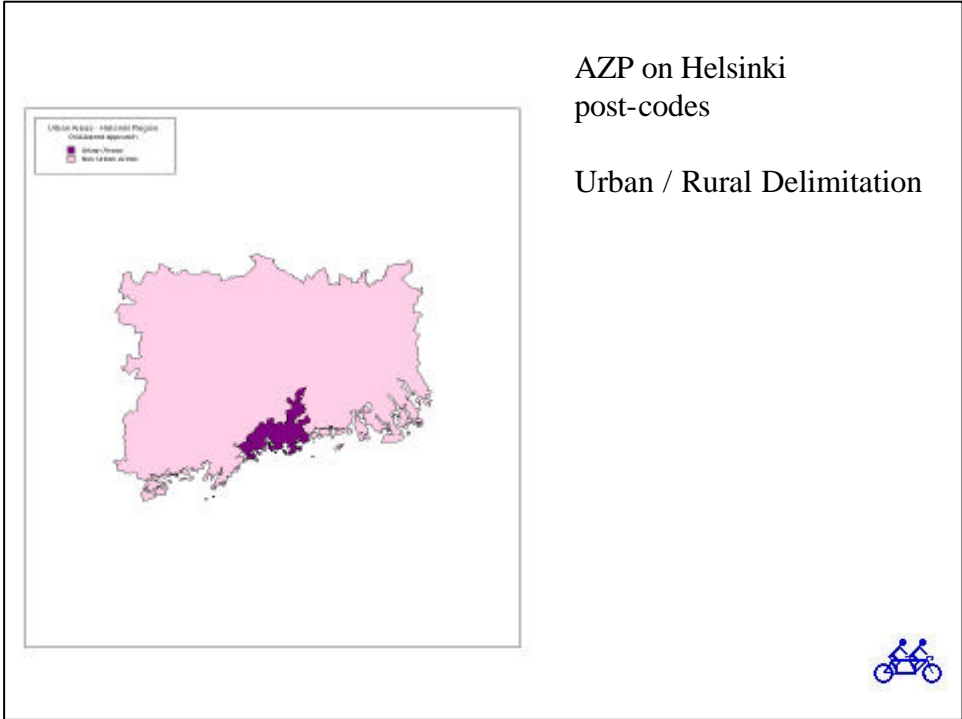
Building Blocks





AZP on South Glamorgan
Enumeration Districts

Urban / Rural Delimitation



AZP on Helsinki
post-codes

Urban / Rural Delimitation



Conclusion of the analyses (1)

Both methods propose their own way to harmonise data production.

- To perform well both approaches:

- ? Need the provision of a 'good' initial set of areal units.

- ? Need to incorporate confidentiality restrictions (for the dissemination of statistics).

- The AZP method incorporates it within the internal process of calculation,

- The grid-based approach incorporate it in a step-by-step process performed by the user;

- ? Face the definition of optimality (e.g. optimal grid-size, optimisation of an objective function);



Conclusion of the analyses (2)

- Need to deal with limitations due to processing speed and disk space;
- A good GIS software environment.

- Both methods did not manage to take into account several specific problems:

- For example how the data near outer boundaries of the study area is taken account (grid-based) or data of islands or other special physical structures (both);

- The AZP algorithm is able to incorporate social constraints. This aspect has not been considered within the framework of this project.



Conclusions of the project

- It was a part of the project to organise the technical implementation of the both grid-based and region-based systems.
- This study is hopefully one step towards a common geographical base to compare statistics across the Europe. There is a need for methodological developments.
- There is also a need for more case studies with different kind of input data, with different scale of analyses and with different type of spatial analyses.



Results and deliverables

- Reports of theoretical and technical assessment of two alternative building blocks for a system of geo-statistics of Europe
- Prototypes of grid (regular)-based and polygon(irregular)-based statistical system
- Results of delineation of urban areas with the building blocks
- Recommendations



The Tandem Consortium